chemistry-climate models of the stratosphere

J. AUSTIN¹, D. SHINDELL², C. BRÜHL³, M. DAMERIS⁴, E. MANZINI⁵, T. NAGASHIMA⁶, P. NEWMAN⁷, S. PAWSON⁷, G. PITARI⁸, E. ROZANOV⁹, C. SCHNADT⁴, T.G. SHEPHERD¹⁰

² NASA-Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025.
³ Max Planck Insitut für Chemie, Mainz, Germany.
⁴ DLR, Oberpfaffenhofen, Weßling, Germany.
⁵ Max Planck Insitut für Meteorologie, Hamburg, Germany.
⁶ Center for Climate System Research, University of Tokyo, Japan.
⁷ Goddard Earth Sciences and Technology Center, NASA/Goddard Space Flight Center Code 916, Greenbelt, MD 20771.
⁸ Dipartamento di Fisica, Università de L'Aquila, 67010 Coppito, L'Aquila, Italy.
⁹ PMOD-WRC/ IAC ETH, Dorfstrasse 33, Davos Dorf CH-7260, Switzerland.
¹⁰ Atmospheric Physics, University of Toronto, Toronto, Ontario, Canada.

December 7, 2001

To be submitted to the Journal of Geophysical Research
In recent years a number of chemistry-climate models have been developed with an emphasis on the stratosphere. Such models cover a wide range of timescales of integration using models of varying complexity. The results of specific diagnostics are here analysed to examine the strengths and weaknesses of the individual models. Many models indicate a significant cold bias in high latitudes, the 'cold pole problem', particularly in the southern hemisphere during winter and spring. This problem can be alleviated with the use of non-orographic gravity wave drag schemes which also considerably improves the relationship between temperature and heat fluxes from the lower atmosphere. The widely varying ozone values are seen to result from the widely differing amounts of polar stratospheric clouds simulated by the models.

The results are also compared to determine the possible future behaviour of ozone, with an emphasis on the polar regions and mid-latitudes. The different models indicate the ozone depletion may worsen but then give a range of results concerning the timing and extent of ozone recovery. Differences in the simulation of gravity waves and planetary waves as well as model resolution are likely major sources of uncertainty for this issue. In the Antarctic, the ozone hole has probably reached almost its deepest although the vertical and horizontal extent of depletion may increase slightly further over the next few years. According to the model results, ozone recovery could begin any year within the range 2001 to 2008. For the Arctic, most models indicate that small ozone losses may continue for a few more years and that recovery could begin any year within the range 2004 to 2019. The start of ozone recovery in the Arctic is therefore expected to appear later than in the Antarctic in most models. Further, interannual variability will tend to mask the signal for longer in the Arctic than in the Antarctic, delaying still further the date at which ozone recovery may be said to have occurred.

1. Introduction

It has been known for some time (e.g. Groves and Tuck, 1980) that increases in greenhouse gases (GHGs) have an impact on ozone amounts in the stratosphere by changing its thermal structure and hence the rate at which ozone is destroyed. Changes in the thermal structure also cause changes in wind fields which affect the transport of ozone and long-lived species which affect ozone chemically. By changing the amount of ozone in the stratosphere, the amount of UV radiation reaching the troposphere is also affected, leading to changes in tropospheric ozone and tropospheric chemistry and climate. Despite considerable research on the topic, the extent to which stratospheric change can influence climate is only beginning to be understood. The discovery of the Antarctic ozone hole (Farman et al., 1985) has added further complexity because whereas previously it was thought that increases in GHGs cool the stratosphere and increase ozone, the processes leading to the ozone hole (e.g. Solomon, 1986) may under some circumstances lead to a decrease in ozone as GHGs increase, because of an increase in Polar Stratospheric Clouds (PSCs). Possibly one of the most extreme examples of chemistry-climate coupling is the effect of increasing GHGs on Arctic ozone. Using a mechanistic model with reasonably comprehensive chemistry Austin et al. (1992) showed that a doubling of $CO_2$ concentrations, expected towards the end of the 21st century, could lead to severe Arctic ozone loss if large halogen abundances persisted until that time. On
in Arctic ozone due to a $CO_2$ doubling, while again keeping chlorine amounts fixed.

Since the early 1990s, the amendments to the Montreal protocol have resulted in a considerable constraint on the evolution of halogen amounts and more recent calculations have been able to take this into consideration. For example, in a coupled chemistry-climate simulation, Shindell et al. (1998) specified currently projected concentrations of halogens and GHGs and calculated increased ozone depletion over the next decade or so, with severe ozone loss in the Arctic in some years. In contrast, recent results from other models (Austin et al., 2000; Brühl et al., 2001; Nagashima et al., 2001; Schnadt et al. 2001; Austin and Butchart, 2001) indicate a relatively small change in ozone over the next few decades. Further, while most models suggest that increases in radiative cooling from GHG increases dominate, giving rise to increased ozone depletion, in reality atmospheric processes may be more complex. For example, Schnadt et al. (2001) predict an increase in ozone over the period 1990-2015 due to GHG increases, because their model simulates an increase in planetary wave activity and an associated warming of the northern polar stratosphere. However, even though a given model may indicate a statistically significant signal, the results from different models may well be different, because of their sensitivity to details such as the position of the subtropical jet (e.g. Hartmann et al., 2000).

This illustrates one of the problems concerning future predictions: the simulation of the north polar stratospheric temperature and thus, ozone, is complicated by the models' ability to reproduce realistically the dynamical activity of this region. The model results therefore crucially depend on the length of the numerical simulation and model configuration, such as spatial resolution, position of the model upper boundary, gravity wave drag scheme and the degree of coupling between the chemically active species and the radiation scheme. The fact that the results are model dependent emphasises that the various dynamical-chemical feedback mechanisms in the atmosphere, as well as the dynamical coupling between troposphere and stratosphere, are not yet fully understood. These mechanisms are here explored by comparing specific diagnostics from a range of coupled chemistry-climate models. In addition to providing increased confidence in the model results, these comparisons can reveal evidence regarding the performance of the individual models. For example, while it might generally be expected that model resolution improves the results, not all factors may be significantly affected or indeed improved. The primary interest here is polar ozone and factors which control it, such as PSCs. Polar ozone also has an impact on middle latitudes via transport, while selected globally averaged quantities are also considered when they offer general insights into overall model performance.

In Section 2, the 3-D coupled chemistry-climate models used in this comparison are described. In Section 3, several factors known to affect model predictions of ozone are investigated using, as indicated above, specific diagnostics designed to illustrate model differences and uncertainties. These diagnostics are compared amongst the models and with observations. Having established model strengths and weaknesses, their results are assessed in Section 4, to try to provide a consensus on the likely trend in future ozone. Conclusions are drawn in Section 5.

2. Models used in the comparison

The models used in the analysis are indicated in Table 1, in order of decreasing horizontal resolution.
Traditionally, climate models have been run with fixed GHGs for both present and doubled CO\textsubscript{2} with the investigation of the subsequent ‘equilibrium climate’. Several coupled chemistry-climate runs have followed this route with multi-year ‘timeslice’ simulations applicable to GHG concentrations for specific years (e.g. Schnadt et al., 2001; Pitari et al. 2001; Rozanov et al., 2001). Other climate simulations have involved transient changes in the GHGs, and several coupled chemistry-climate simulations have followed this pattern (e.g. Shindell et al., 1998a; Austin 2001; Nagashima et al., 2001). The advantage of transient experiments is that the detailed evolution of ozone can be determined in the same way that it is likely to occur (in principle) in the atmosphere, albeit with some statistical error. The impact of interannual variability can be assessed by determining this variability over, e.g. a 10-year period. Comparisons with observations are direct since both atmosphere and model cover the period when GHGs and halogens are changing to the same extent. Timeslice simulations need a sufficient duration (at least 20 years) to allow the interannual variability to be determined but in principle, 20 years evaluations of the same conditions may have less variability than the atmosphere in which halogens may be changing rapidly. Timeslice runs also have the advantage that several realisations of the same year are available, from which future predictions can be assessed. However, in practice this may be of less value than examining the behavior of different models, since a given model will tend to have systematic errors. Both transient simulations and timeslice simulations are here used, to bring together the best of both sets of simulations.

**UMETRAC**

The Unified Model with Eulerian TRansport And Chemistry is based on the Met. Office's Unified Model, which has been used in the Intergovernmental Panel on Climate Change (IPCC) assessments (e.g. IPCC, 2001). UMETRAC has 64 vertical levels, from the ground to 0.01 hPa, and includes a coupled stratospheric chemistry scheme involving 13 advected tracers ($O_3$, $HNO_3$, $N_2O_5$, $HNO_4$, $HCl$, $HOCl$, $ClONO_2$, $HOBr$, $HBr$, $BrONO_2$, $H_2O_2$, $H_2CO$, $Tracer-1$). $Tracer-1$ is used to parameterise the long-lived species $H_2O$, $CH_4$, $Cl_y$, $Br_y$, $H_2SO_4$ and $NO_y$. A simplified sedimentation scheme is included and the PSC scheme is based on liquid ternary solutions (LTS) and ice or nitric acid trihydrate (NAT) and ice. The photochemistry therefore explicitly includes all the major processes affecting stratospheric ozone.
surface temperatures and sea ice are specified from observations or from simulations of a coupled ocean-atmosphere version of the model. The dynamics is available using Rayleigh friction as standard to decelerate the jet or with a non-orographic gravity wave drag (gwd) scheme. The latest results of the model are described in Austin (2001) (Rayleigh Friction) and Austin and Butchart (2001) (non-orographic gwd scheme).

MA-ECHAM CHEM

**ECHAM4.L39(DLR)/CHEM**

A detailed description of the chemistry-climate model ECHAM4.L39(DLR)/CHEM (hereafter E39/C) has been given by Hein et al. (2001), which also discussed the main features of the model climatology. The model horizontal resolution is T30 with a corresponding Gaussian transform latitude-longitude grid, on which model physics, chemistry, and tracer transport are calculated, with mesh size $3.75^\circ \times 3.75^\circ$. In the vertical, the model has 39 layers (L39) from the surface to the top layer centered at 10 hPa (Land et al. 1999). A parameterization for orographic gravity wave drag (Miller et al., 1989) is employed, but the effects of non-orographic gravity waves are not considered. The chemistry module CHEM (Steil et al. 1998, updated Hein et al., 2001) is based on the family concept, containing the most relevant chemical compounds and reactions necessary to simulate upper tropospheric and lower stratospheric ozone chemistry, including heterogeneous chemical reactions on polar stratospheric clouds (PSCs) and sulphate aerosol, as well as tropospheric $\text{NO}_2-\text{HO}_2-\text{CO}-\text{CH}_4-O_3$ chemistry. Physical, chemical, and transport processes are calculated simultaneously at each time step, which is fixed to 30 minutes. The PSCs in E39/C are based on NAT and ICE amounts depending on the temperature and the concentration of $\text{HNO}_3$ and $\text{H}_2\text{O}$ (Hanson and Mauersberger, 1988). A nucleation barrier is prescribed to account for the observed super-saturation of NAT-particles (Schlager et al., 1990; Peter et al., 1991; Dye et al., 1992), A simplified scheme for the sedimentation of NAT and ice particles is applied. Stratospheric sulphuric acid aerosol surface areas are based on background conditions (WMO 1992, Chapter 3) with a coarse zonal average. Sea surface temperature and sea ice distributions are prescribed for the various time slices according to the transient climate change simulations of Roeckner et al. (1999).

CCSR/NIES

**UIUC** (University of Illinois at Urbana-Champaign) is a global-scale grid-point General Circulation Model with interactive chemistry (Yang et al., 2000; Rozanov et al., 2001). The model horizontal resolution is $4^\circ$ latitude and $5^\circ$ longitude, with 24 layers spanning the atmosphere from the surface to 1 hPa. The chemical-transport part of the model simulates the time-dependent three-dimensional distributions of 42 chemical species ($\text{O}_3$, $\text{O}^{(1)}\text{D}$, $\text{O}^{(3)}\text{P}$), $\text{N}$, $\text{NO}$, $\text{NO}_2$, $\text{NO}_3$, $\text{N}_2\text{O}_5$, $\text{HNO}_3$, $\text{HNO}_4$, $\text{N}_2\text{O}$, $\text{H}$, $\text{OH}$, $\text{HO}_2$, $\text{H}_2\text{O}_2$, $\text{H}_2\text{O}$, $\text{H}_2$, $\text{Cl}$, $\text{ClO}$, $\text{HCl}$, $\text{HOCl}$, $\text{ClNO}_3$, $\text{Cl}_2$, $\text{Cl}_2\text{O}_2$, $\text{CF}_2\text{Cl}_2$, $\text{CFCl}_3$, $\text{Br}$, $\text{BrO}$, $\text{BrNO}_3$, $\text{HOBr}$, $\text{HBr}$, $\text{BrCl}$, $\text{CBrF}_3$, $\text{CH}_3\text{Br}$, $\text{CO}$, $\text{CH}_4$, $\text{CH}_3\text{O}_2$, $\text{CH}_3\text{OOH}$, $\text{CH}_3\text{O}$, $\text{CH}_2\text{O}$, and $\text{CHO}$), which are determined by 199 gas-phase and photolysis reactions. The model also takes into account 6 heterogeneous reactions on and in sulfate aerosol and polar stratospheric cloud particles. The chemical solver is based on the pure implicit iterative Newton-Raphson scheme.
and Sander et al. (2000). Photolysis rates are calculated at every chemical step using a look-up-table (Rozanov et al., 1999). The advective transport of species is calculated using the hybrid advection scheme proposed by Zubov et al. (1999). A simplified scheme is applied for the calculation of the PSC particles (NAT and ice) formation and sedimentation. Sea surface temperature and sea ice distributions are prescribed from the AMIP-II monthly mean distributions, which are the averages from 1979 through 1996 (Gleckler, 1996). The orographic scheme of Palmer et al. (1986) is used to parameterize gravity wave drag. In this paper the results of a 15-year long control run of the model will be considered. The climatology of this run has been presented by Rozanov et al. (2001).

ULAQ
GISS

3. Model uncertainties
3.1 The Cold Pole problem

Many climate models without chemistry but with a fully resolved stratosphere have a cold bias of the order of 5-10K over the South Pole in the lower stratosphere, suggesting that the residual circulation is too weak (Pawson et al., 2000), i.e. there is too little downwelling in high latitudes or too little upwelling in low latitudes. This temperature bias could have a significant impact on model heterogeneous chemistry, and enhance ozone destruction. The 'cold pole problem' extends to higher levels in the stratosphere and by the thermal wind relation gives rise typically to a polar night jet that is too strong and which has an axis which is vertically oriented, whereas in the observations the polar night jet axis slopes with height towards the equator in the upper stratosphere. The weaker jet and slope of the axis allow waves to propagate into higher latitudes and maintain higher temperatures. A potentially important component of climate change is whether these waves increase in amplitude with time since this will likely affect the evolution of ozone: see section 3.5. The practical solution to those models with a cold bias is to adjust the temperatures in the heterogeneous chemistry (e.g. Austin et al., 2000) to prevent extensive ozone depletion from occurring. The strong polar night jet is also associated with a south polar vortex that breaks down later in the spring, particularly in the southern hemisphere. In a coupled chemistry-climate model this can lead to ozone depletion that continues for longer than observed and hence produces more complete destruction in the ozone hole, with a longer lasting ozone hole.

Recently the development of non-orographic gravity wave drag (gwd) schemes for climate models (e.g Hines, 1997; Medvedev et al., 1998; Warner and McIntyre, 1999) has resulted in a significant reduction in the cold pole problem relative to simulations which rely on Rayleigh friction to decelerate the polar night jet (e.g. Manzini and McFarlane, 1998). Two of these schemes have also been shown to produce a quasi-biennial oscillation (QBO) when run in a climate model (Scaife et al., 2000a). The latest versions of several coupled chemistry-climate models now employ such schemes: CMAM uses the Medvedev-Klaassen scheme (Medvedev et al. 1998) or the Hines scheme; UMTRAC uses Warner and McIntyre (1999) (Austin and Butchart, 2001); and MA-ECHAM4-CHEM uses Hines et al. (1997) (Steil et al., 2001). The GISS GCM has used a non-orographic gravity wave drag scheme for many
Figure 1: Temperature biases at 80°N and 80°S for the winter and spring seasons, as a function of pressure. To determine the bias, a climatology determined from 10 years of UKMO data assimilation temperatures were subtracted from the model results.
reasonably well (Shindell et al., 1998b) but does not simulate a QBO in the tropics.

Figure 1 shows model temperature biases as a function of height for 80°N and 80°S for the winter and spring seasons. To determine the biases, a 10-year temperature climatology determined from data assimilation fields (Swinbank and O'Neill, 1994) was subtracted from the mean model temperature profiles applicable to the 1990s. The upper stratospheric cold pole problem is particularly noticeable in the south in the UMETRAC (with Rayleigh friction), CCSR/NIES (which also uses Rayleigh friction) and MAECHAM CHEM results. In the results of the UIUC and ULAQ models a significant cold bias is not present, possibly because the close proximity of the upper boundary, combined with relatively low vertical resolution. Indeed, these models have a warm bias in the middle stratosphere in some months in high southern latitudes. As seen in the UMETRAC results, the winter and spring cold pole can be dramatically relieved by the use of non-orographic gwd. Although some temperature bias remains in the MAECHAM CHEM model, which uses the Hines non-orographic gwd scheme, the results are an improvement on the ECHAM4.L39(DLR)/CHEM which is a very similar model, but does not have a non-orographic gwd scheme. At 80°N temperature biases are somewhat smaller than at 80°S and are sometimes positive. The northern lower stratospheric temperature biases would generally lead to insufficient heterogeneous ozone depletion in early winter but excessive ozone depletion in the more important spring period. In the southern hemisphere, spring cold biases could lead to more extensive PSCs than observed and delayed recovery in Antarctic ozone.

Figure 2: Scatter diagrams of heat flux $\overline{\nu T'}$ (averaged 40° - 80°N, at 100 hPa for Jan and Feb) against temperature (averaged 60° - 90°N, at 50 hPa for Feb and March) for participating models. The solid lines are linear regression lines between the two variables. The observations are taken from NCEP data. The heat flux is in units of Kms$^{-1}$, the temperature is in K.
Table 2. Statistical analysis of the linear regression between the area averaged temperature (K) at 50 hPa polewards of 60°N for Feb and March, and the heat flux (Kms⁻¹) at 100 hPa between 40° and 80°N for Jan. and Feb (northern hemisphere). The southern hemisphere results are for the months Aug. and Sept. and July and Aug. respectively. $R^2$ is the correlation coefficient between the variables, $T_0$ is the intercept of the line at zero heat flux, and $\beta$ is the gradient of the line.

<table>
<thead>
<tr>
<th>Model/Observations</th>
<th>N’ern Hemi.</th>
<th>S’ern Hemi.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>$T_0$</td>
</tr>
<tr>
<td>NCEP (Observations)</td>
<td>0.77</td>
<td>195.1</td>
</tr>
<tr>
<td>UMETRAC Non-orographic gwd</td>
<td>0.74</td>
<td>196.9</td>
</tr>
<tr>
<td>Rayleigh Friction</td>
<td>0.67</td>
<td>196.2</td>
</tr>
<tr>
<td>ECHAM/CHEM 1960</td>
<td>0.79</td>
<td>196.3</td>
</tr>
<tr>
<td>ECHAM4.L39(DLR)/CHEM</td>
<td>0.83</td>
<td>190.4</td>
</tr>
<tr>
<td>CCSR/NIES</td>
<td>0.62</td>
<td>198.3</td>
</tr>
<tr>
<td>ULAQ</td>
<td>0.74</td>
<td>199.4</td>
</tr>
</tbody>
</table>

Some indication of the source of the model temperature biases in the lower stratosphere is given by Figure 2, which shows the heat flux $\overline{v'T}$ at 100 hPa averaged over the domain 40°-80°N for January and February plotted against temperature averaged over the domain 60°-90°N at 50 hPa for February and March. As argued by Newman et al. (2001), the heat flux at 100 hPa is indicative of the wave forcing from the troposphere and this is highly correlated with lower stratospheric temperature slightly later in the year, once the waves have propagated. Newman et al. choose the period of the temperature average as 1-15 March, which max-
we choose a longer period for the temperature average to smooth over any model and atmospheric transients. This reduces the correlation coefficient for NCEP data only slightly, to 0.77 (0.78 in the south. The model results (Figures 2 and 3) are in general agreement with observations and indicate a range in the correlation coefficient of 0.58 to 0.83 in the North and 0.51 to 0.86 in the South (Table 2). The results indicate that horizontal resolution may have affected the model results: in general the gradients of the model regression lines (β in Table 2) become less steep as the model resolution decreases, particularly in the northern hemisphere. Thus models are better able to capture the low-amplitude wave, small heat flux case than the large heat flux case with its potentially significant cascade to larger wavenumbers. The exception to this general rule is the ECHAM/CHEM model simulation for 1990 which has a large value of β in good agreement with observations. Extrapolating the regression line to zero flux indicates the radiative equilibrium temperature $T_0$ (e.g. Newman et al., 2001). For the ECHAM/CHEM 1990 model simulation this is somewhat lower than for the other simulations. In the southern hemisphere $T_0$ is considerably lower, due to lower ozone amounts, and all the models agree well with the observed $T_0$. For the timeslice experiments (e.g. MA-ECHAM CHEM) $T_0$ decreased with time due to the GHG increase. The values of β are also generally much smaller in the southern hemisphere, except for the CCSR/NIES model. There is also an indication from the UMETRAC results that non-orographic gwd can significantly improve the temperature-heat flux relationship although this would need to be confirmed by other models. In contrast non-orographic gwd has a much smaller impact on the results in the northern hemisphere.

3.2 The simulation of polar stratospheric clouds

During the last few years, considerable progress has been made regarding the understanding of PSCs and the associated heterogeneous reactions. Some of these developments are discussed in Carslaw et al. (2001) and WMO (2002), Chapter 3. They include the observation of many different types of PSCs, both liquid and solid forms and the laboratory measurement of a wide range of chemical reactivities as summarised in Sander et al. (2000). Consequently, models have a variety of PSC schemes with and without sedimentation. So that common ground can be established between models, we here use the temperature at 50 hPa as an indicator, hence ignoring the impact of HNO$_3$ and or sulphate concentrations on the determination of PSC surface areas.

While models have significant temperature biases in their climatologies, as indicated in Section 3.1 a more relevant issue for polar ozone depletion is the behavior of climate models near PSC temperatures. Following Pawson and Naujokat (1997) and Pawson et al. (1999), Figure 4 shows the time integral of the winter PSCs in each of the participating models in comparison with observations. This is the quantity $\bar{A}_r$ indicated in Figure 9 of Pawson et al. (1999), but without the normalisation factor of 1/150, for $\tau_{\text{NAT}} = 195$K and $\tau_{\text{ICE}} = 188$K. The areas in each of the models varies dramatically, particularly for the ice amount. In the Arctic, this varies between zero (ULAQ model, not shown) and 70 in units of % of the hemisphere times days for the mean with large interannual standard deviation in the timeslice runs. Arctic NAT also covers a large range, both for different models and in the interannual variability for each simulation. In the Antarctic each model has much lower fractional interannual variation, but again the results for separate models cover an exceedingly large
Figure 4: Seasonal integration of the hemispheric area at 50 hPa, corresponding to the temperatures 195K (approximate NAT temperature) and 188K (approximate ICE temperature) for participating models. The error bars for the timeslice runs indicate 95% confidence intervals. For clarity, ULAQ data are plotted two years early, and for 1990 and 2030 MAECHEM CHEM data are plotted two years late.
on the amount of ozone depletion calculated.

The sedimentation of PSC particles is now thought to be important in polar ozone depletion (e.g. Waibel et al., 1999), but is absent from many coupled chemistry-climate models. Using a box model, Waibel et al. demonstrated that cooling associated with greenhouse gas increases may lead to higher ozone depletion than if sedimentation were ignored. There are many practical difficulties associated with incorporating sedimentation in climate models. The full details, including dividing the particles into different sized bins (e.g. Timmreck and Graf, 2000) cannot be readily simulated as this would require the transport of particles covering a large number of size ranges, which may exceed the total number of tracers transported. Most models therefore adopt some extremely simplified procedures. For example in the UMETRAC model (Austin, 2001) sedimentation is assumed to occur for NAT or ice particles of fixed size. The vertical transport equation is solved locally to give a sedimented field which is not explicitly advected horizontally but is tied to the field of a conserved horizontal tracer. This is necessary because of the treatment of long-lived tracers such as $H_2O$ and $NO_y$ which are also tied to the same horizontal tracer. Other models, which transport $H_2O$ and $NO_y$ explicitly (e.g. Egorova et al., 2001; Hein et al., 2001, some more refs) are able to transport the sedimented field in more detail, but the accuracy of this transport may be limited by the accuracy of the transport scheme, and whether it can treat small scale features with high enough precision. Recently, the presence of large, nitric acid containing particles exceeding 10 $\mu m$ in diameter has been detected (Fahey et al., 2001) in the lower stratosphere. However, models do not yet take the production of these particles into account, and yet they may have an important impact on ozone depletion.

3.3 The position of the upper boundary

There is strong evidence from a number of modeling studies (Garcia and Boville, 1994; Shepherd et al., 1996; Lawrence, 1997; Austin et al., 1997; Rind et al., 1998; Beagley et al., 2000) that the position of the model upper boundary can play a significant role in influencing transport and stratospheric dynamics due to the ‘downward control principle’ (Haynes et al., 1991). The sensitivity to the position of the upper boundary may be more when using non-orographic gwd schemes than when Rayleigh friction is used, although if all the non-orographic gwd that is produced above the model boundary is placed instead in the top model layer, assuming that it is not a sponge layer, this sensitivity reduces (Lawrence, 1997). Model simulations with an upper boundary as low as 10 hPa have been completed (e.g. Schnadt et al., 2001; Hein et al., 2001; Dameris et al., 1998). Schnadt et al. show the meridional circulation of the DLR model and this gives the expected upward motion from the summer hemisphere and downward motion over the winter hemisphere. Their argument is that it is less important to have a high upper boundary, but more important to have high resolution in the vicinity of the tropopause. However, there is evidence, for example in the total ozone presented by Hein et al., that in the same model insufficient ozone is transported to the north pole. Also, modelled meridional circulations are known to extend into the mesosphere (e.g. Butchart and Austin, 1998). While the transport effect on ozone is direct, other considerations are the transport of long-lived tracers such as nitrogen oxides and water vapor which have a photochemical impact on ozone. Consequently, it is generally recognised that the upper boundary should be placed at least as high as 1 hPa (e.g. Rozanov et al.,
0.01 hPa (e.g. Shindell et al., 1998; Austin et al., 2001; Steil et al., 2001, Nagashima et al., 2001). In comparison, the Canadian Middle Atmosphere Model (de Grandpré et al., 2000) has an upper boundary somewhat higher (c. 0.0006 hPa) to allow a more complete representation of gwd to reduce the cold pole problem (Section 3.1) and to simulate upper atmosphere phenomena.

The impact of the upper boundary can be assessed by investigating the Brewer-Dobson calculation computed for the individual models. An indicator of the circulation is given by the residual mean meridional circulation

$$\mathbf{v}^* = \mathbf{v} - \frac{\partial}{\partial p} \left( \frac{\mathbf{v} \cdot \mathbf{\theta}'}{\partial \theta / \partial p} \right), \quad \mathbf{\omega}^* = \mathbf{\omega} + \frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} \left( \cos \phi \frac{\mathbf{v} \cdot \mathbf{\theta}'}{\partial \theta / \partial p} \right)$$

For details of the notation see Edmon et al., 1980. The residual circulation ($\mathbf{v}^*, \mathbf{\omega}^*$) is then used to compute the residual streamfunction $\Psi$, defined by

$$\frac{\partial \Psi}{\partial \phi} = -a \cos \phi \mathbf{\omega}^*, \quad \frac{\partial \Psi}{\partial p} = \cos \phi \mathbf{v}^*$$

Figure 5 shows the mass streamfunction $F_m = 2\pi a \Psi / g$ at 50 hPa as a function of latitude for the participating models.

### 3.4 Interannual variability

In the Arctic, the atmosphere shows considerable interannual variability associated with planetary wave fluxes from the troposphere. By examining almost two decades of satellite data, Scaife et al. (2000b) concluded that at least 10 years of observations were required to account for most of the interannual variability. Consequently, climate model simulations need also to be at least 10 years (and preferably at least 15) to capture model climatology and to be able to determine any trends. In practice, some models have larger variability than observed and this has a direct impact on ozone (e.g. UMETRAC results in the southern hemisphere when run with a non-orographic gwd scheme Austin and Butchart, 2001).

Interannual variability is also significant in globally averaged quantities such as temperature. This variability can be enhanced by incorporating coupled chemistry into models. Figure 6 shows the interannual standard deviation of the temperature from the trend for SSU observations (Scaife et al., 2000b) in comparison with model results. CCSR/NIES, UIUC and UMETRAC results compare favourably with observations in this respect, but in the stratosphere, the Met. Office Unified Model (similar to UMETRAC but without interactive chemistry) has much reduced interannual variability. However, the MA ECHAM CHEM results (with fixed GHGs) indicate low interannual variability, despite the use of coupled chemistry.

### 3.5 Future predictions of planetary waves

In some GCMs, there is a significant trend in planetary wave propagation with time. In the GISS GCM, planetary waves are refracted equatorward as greenhouse gases increase (Shindell et al., 2001) while in the ULAQ model a marked reduction in the propagation of planetary waves 1 and 2 to high northern latitudes in the doubled CO₂ climate is simulated by Pitari et al. (2001). In the GISS model the...
Figure 5: Residual streamfunction ($10^9$ Kgs$^{-1}$) at 50 hPa as a function of latitude and season for the participating models.
The direct radiative cooling by greenhouse gases at high latitudes in the lower stratosphere causes an increase in the strength of the polar vortex. Planetary wave changes may be a feedback which strengthens this effect. One proposed planetary wave feedback mechanism (Shindell, 2001) works as follows: (1) tropical and subtropical SSTs increase, leading to (2) a warmer tropical and subtropical upper troposphere via moist convective processes. This results in (3) an increased latitudinal temperature gradient at around 100-200 hPa, as these pressures are in the stratosphere at higher latitudes, so do not respond to the surface warming. The temperature gradient leads to (4) enhanced lower stratospheric westerly winds, which (5) refract upward propagating tropospheric planetary waves equatorward. This results in (6) a strengthened polar vortex, as the reduced upward propagation at high latitudes limits the ability of individual waves to enhance abruptly the residual circulation and cause sudden warmings.

However, climate experiments containing different physics with higher spatial resolution models (e.g. Schnadt et al., 2001) do not show a future trend towards re-
personal communication, 2001) predicts a future increase in overall generation of planetary waves, which leads to a greater wave flux to the Arctic stratosphere, and is even able to overcome the radiatively induced increase in the westerly zonal wind so that the overall trend is to more easterly flow. This also occurs in the DLR model with chemical feedback (Schnadt et al., 2001) whereas in the UM with chemical feedback the downward trend in the heat flux during the period 1975-2020 is not statistically significant (Figure 7). The NIES model, of lower resolution than UMETRAC, has systematically lower heat fluxes but does show a downward trend during the period 1986-2050 which is marginally statistically significant. In general, the strengthening of the polar vortex appears to be critically dependent upon the relative importance of changes in wave generation versus wave propagation. These likely are highly model and resolution dependent, resulting from the particular wave forcing and drag schemes employed in each climate model.

Observations suggest that in the past twenty years the Arctic vortex has strengthened (Tanaka et al., 1996; Zurek et al., 1996; Waugh et al., 1999; Hood et al., 1999). Likewise, a trend toward equatorward wave fluxes and downward propagation of wind anomalies has also occurred in observations (Kodera and Koide, 1997; Kuroda and Kodera, 1999; Ohhashi and Yamazaki, 1999; Baldwin and Dunkerton, 1999, Hartmann et al., 2000). A comparison of the zonal wind and wave flux anomalies seen in the GISS GCM and in NCEP data is shown in Shindell et al. (2001), Plate 3. Other model simulations show qualitatively similar effects on planetary wave propagation (Kodera et al., 1996; Perlwitz et al., 2000). However, current trends in the observations in the Arctic have been affected by some recent warm years, and hence extrapolation of those trends to the future must remain uncertain.

Figure 7: Scatter diagrams of heat flux $\vec{V}T$ (averaged 40° - 80°N, at 100 hPa for Jan and Feb) against year for participating models. Linear regression lines between the two variables are also drawn. For UMETRAC, the dashed lined is for the non-orographic gwd run, and the dot-dash line is for the Rayleigh friction run. The observations are taken from NCEP data.
Seven years of global satellite observations of the middle and upper stratosphere, supported by coincident ground-based data (Nedoluha et al., 1998; Randel et al., 1999), and lower stratospheric measurements over Boulder, Colorado and Washington, DC, covering a period of 37 years (Oltmans and Hofmann, 1995; Oltmans et al., 2000) all show very large increases of water vapor, in the range of 8-20% per decade. A summary analysis of ten data sets taken by five types of instruments (Halogen Occultation Experiment (HALOE) satellite, balloon, aircraft, and ground-based) which have reported observations of stratospheric water vapor during 1954-2000, shows an overall increase of 1% yr\(^{-1}\), (Rosenlof et al., 2001). Some aircraft observations, however, give conflicting indications of trends in the lower stratosphere (Peter, 1998, Hurst et al., 1999). The evidence largely supports a significant trend in the past, suggesting that a similar trend may take place over the coming decades.

The calculated enhanced rate of water vapor production due to methane increases is sufficient to account for only about 30% of the observed water vapor increase. It is therefore suspected that there has been increased transport from the troposphere to the stratosphere. This is governed largely by tropical tropopause temperatures (e.g. Mote et al., 1996), and could be greatly altered by changes as small as a few tenths of a degree (Evans et al., 1998). Observations do not indicate an overall warming trend at the tropical tropopause (Simmons et al., 1999). However, water vapor fluxes as a function of season and geographic location have not been precisely quantified (SPARC, 2000), so that it is difficult to determine what may have in fact caused an increased flux to the stratosphere. For example, a very small shift away from the equator could raise the minimum temperatures encountered sufficiently to allow a significantly greater upward flux. Cross-tropopause transport can also be limited by water vapor availability, so that an increase in tropospheric water vapor, predicted by most climate models, would therefore lead to increased stratospheric water even without a warmer tropopause. Recently, Hartmann et al. (2001) have described a possible mechanism whereby lowering tropopause temperatures can lead to increased flux of water into the stratosphere. This results from the increased formation and transport of ice particles into the lower stratosphere.

Increased water vapor affects both ozone chemistry directly, and also alters local temperatures by radiative cooling, slowing down the reaction rates of ozone depletion chemistry, and therefore indirectly leading to more ozone. The effects on homogeneous chemistry have been studied by Evans et al (1998), Dvortsov and Solomon (2001), and Shindell (2001). The models all show that increases in water vapor lead to reduced ozone in the upper stratosphere, increased ozone in the middle stratosphere, and reductions in the lower stratosphere. The latter dominates the total column, and the model results differ most in this region. In the model of Evans et al. (1998), lower stratospheric ozone is reduced only in the tropics when water vapor increases, while in the others, reductions extend to mid-latitudes or the poles. Thus the models of Dvortsov and Solomon (2001) and Shindell (2001) show a slower ozone recovery by about 10-20 years, and a 1-2% reduction during the next 50 years due to water vapor increase. The Evans et al. (1998) model disagrees with these results, presumably due to differences in the model’s temperature response to increasing water, which seems to dominate over its chemical impacts in that model.

Water vapor also affects heterogeneous chemistry, enhancing the formation of PSCs. This may be much more important than the relatively small impacts of water on homogeneous chemistry. These effects have been studied by Kirk-Davidoff et
significant enhancement to Arctic ozone loss in a more humid atmosphere. Much of this effect is based on the radiative cooling assumed to be induced by water vapor. A relatively large value is used, -6 to -9 K per ppmv of water vapor based on Forster and Shine (1999). However, recent modeling suggests that this value may be smaller, about -1.5 to -2.5 K per ppmv of water (Oinas et al., 2001). This would in turn imply a reduced role for water vapor in enhancing PSC formation. Even using the smaller cooling rate, however, the impact on ozone may be large, as the 3K cooling due to doubled CO₂ is of comparable magnitude to a water vapor increase of only a couple ppmv. Tabazadeh et al (2000) showed that the enhancement of PSC formation due to the addition of 1 ppmv of water vapor is approximately the same as the PSC enhancement due to cooling of about 1 K. This suggest that the radiative impact of water vapor is larger than its effects on chemistry or microphysics. Given the potential for denitrification in the Arctic, and the large ozone losses that could result from a slight cooling there (Tabazadeh et al., 2000), it is important both to understand trends in stratospheric water vapor, and to resolve model differences in the radiative impact of those trends.

Model simulations of past water vapor trends do not agree well with observations. In UMETRAC (Austin, 2001), water vapor increases by only about 1% per decade in the stratosphere, despite the inclusion of a methane oxidation scheme. In UMETRAC the tropical tropopause temperature decreases slightly, counteracting the methane impact. In ECHAM4.L39(DLR)/CHEM (Schnadt et al., 2001), water vapor increases in the lower stratosphere are significantly larger but are still about a factor of 3 lower than observed. The GISS model also reproduces observed increases to some extent (Shindell, 2001). Thus in the models the water vapor trend tends to be driven by the cold trap mechanism, although the Hartmann et al. (2001) mechanism may be equally important in the atmosphere.

4. Model Assessments
4.1 The 1990 - 2020 time frame

As is well established from observations (e.g. WMO, 1999, Chapter 4), ozone has been decreasing over the first part of this time frame, and this is expected to be followed by the first signs of ozone recovery (Shindell et al., 1998a; Austin et al., 2000; Schnadt et al. 2001; Rosenfield et al., 2001; Nagashima et al., 2001; Austin and Butchart, 2001). Two-dimensional model simulations (e.g. Rosenfield et al., 2001) indicate a slight delay in Arctic and Antarctic spring ozone following the maximum values in halogen loading. This may be considered as the ‘radiative effect on chemistry’: GHGs cool the stratosphere and enhance heterogeneous chemistry in the lower stratosphere (while reducing ozone depletion in the middle and upper stratosphere). In the GISS model (Shindell et al., 1998a) there is a further dynamical feedback (described in Section 3.3) involving a reduction of planetary waves which result in a considerably increased signal and severe Arctic ozone loss in the decade 2010 - 2020.

Figure 8 shows the minimum Arctic and Antarctic ozone for a range of coupled chemistry-climate models in comparison with TOMS data. In the Arctic each model has a large interannual variability, similar to that of the observations, and hence detecting a signal is difficult. The continuous lines indicate the 10 year running means of the individual datasets which help to identify the timing of the minima. In the Arctic, all three transient models are in reasonable agreement with observations
than the other models with the simulation indicating a minimum in the smoothed results of about 175 DU compared with almost 100 DU higher in the other transient runs. The date of minimum Arctic ozone, again as indicated by the minimum of the smoothed curves, varies from 2004 for the CCSR/NIES model to 2019 for the GISS model. UMETRAC indicates a minimum at about the year 2015, but the simulation ends shortly afterwards and the smoothed curve is virtually flat in the final decade. All three runs indicate some delay in the onset of ozone recovery, due to increases in GHGs, although such a result is subject to considerable uncertainty because of the large interannual variability.

In the Antarctic, again the transient runs all agree reasonably well with observations for the past and moreover are in fairly good agreement for the future. As in the Arctic, the NIES model indicates the earliest start of ozone recovery (2001) followed by UMETRAC (2005) and GISS (2008). The minima in the smoothed curves are all comparable (109, 86 and 98 DU respectively). This would appear to indicate that ozone recovery will begin earlier in the Antarctic than in the Arctic. Such a quicker recovery, would also be detectable earlier in observations in Antarctica, because of the smaller interannual variability.

The timeslice experiments do not yet have the temporal resolution to give a precise indication of the future timing of ozone recovery. Nonetheless the ECHAM4.L39 /CHEM model (Schnadt et al., 2001) may go against the above picture by suggesting that increases in planetary waves occur in the Arctic speeding up ozone recovery. This may be considered the ‘dynamical effect on chemistry’: Increases in planetary waves transport more ozone as well as raise temperatures and decrease heterogeneous chemistry. Therefore, the net effect on ozone is that of the two potentially competing processes of dynamics and radiation. If planetary waves increase the ‘dynamical’ effect increases ozone and the ‘radiative’ effect decreases ozone, giving a relatively small response. If planetary waves decrease, both the ‘dynamical’ and ‘radiative’ effect are negative, leading to enhanced ozone depletion. To resolve whether increases in GHGs are delaying the onset of ozone recovery, more timeslice simulations are required for the period 1990 to 2015.

The maximum size of the Antarctic ozone hole during each spring, as given by the area within the 220 DU total ozone contour, is shown in Figure 9. The TOMS data are significantly higher than the results from UMETRAC, but recent model runs (H. Struthers, personal communication, 2001), with a revised NOy distribution show an ozone hole about 50% larger for 1995, which is in better agreement with observations. In the observations, the ozone hole does not appear to have reached its maximum by October 2001, indicating the possibility of delayed recovery in the ozone hole. Indeed, the smoothed curve for UMETRAC has its maximum value in 2019, although the curve is virtually flat from 1995 onwards. [Need more model results.....]

The minimum spring ozone columns in mid-latitudes are indicated in Figure 10. These show a similar behavior as for the Arctic and Antarctic with the lowest values occurring in 2019 (Northern hemisphere) and 1999 (Southern hemisphere). However, in the Southern hemisphere, the mean curve is flat from about the year 1995 onwards and hence the timing of ozone recovery is very dependent on the method used to diagnose that recovery. [Need more model results.....]

Possibly of more significance than climate change is the possibility of a future major volcanic eruption. Observations of mid-latitude ozone decreased following the eruption of Mt. Pinatubo (WMO, 1999, Chapter 4) and heterogeneous chemistry is
Figure 8: Minimum Arctic (March/April) and Antarctic (September/October/November) total ozone for participating models. TOMS data are used for comparison. The solid lines indicate the running decadal average for the transient model results and TOMS observations. The broken line is a 4th order polynomial fitted between the UMETRAC snapshot model results. The results for MA-ECHAM-CHEM indicate the mean and range of minimum values obtained in 20 year timeslice experiments for a model output frequency of once every ten days. The model dump frequency may have underestimated slightly the full range of values in the model. To compute the columns, a standard tropospheric column of 40 DU (southern hemisphere) and 100 DU (northern hemisphere) have been added to the stratospheric columns. The results for ECHAM4.L39(DLR)/CHEM indicate the mean and 2 standard deviations for 20 year timeslice runs. The results for ULAQ are for one year integrations of the model from equilibrium conditions.
Figure 9: Maximum area of the Antarctic ozone hole, as given by the 220 DU contour for observations and participating models. See Figure 8 for line styles and symbols.

Figure 10: Minimum total column ozone in the latitude range 30° to 60° in the Northern (left) and Southern (right) hemispheres for participating models. See Figure 8 for line styles and symbols.
canic aerosol to high latitudes could also supply sites for heterogeneous chemistry and contribute to polar ozone depletion. An analysis by Roscoe (2001), based on the frequency of past eruptions, suggests that there is a probability of almost 60% that another major volcanic eruption will occur during the next 50 years. If this were to occur within the next few decades, while halogen levels still remain high, the possibility of large ozone depletions also remains high.

4.2 The 2020 - 2050 time frame

Those models that have run beyond the year 2020 indicate some recovery in ozone. Of particular importance is the return to '1980-like conditions', when anthropogenic halogen concentrations were negligible. As noted in WMO (1999), Chapter 12, this recovery would be to a different vertical distribution of ozone, with higher middle and upper stratospheric ozone. Using a 2-D model, Rosenfield et al. (2001), determined the date for the recovery of total ozone to 1980 levels as a function of day of year and latitude. In the Arctic, this recovery was latest at the end of spring (after 2050) and earliest in Autumn (before 2035). Further, the impact of CO₂ increases was to accelerate the recovery due to increased downwelling. Over Antarctica downwelling has a smaller impact and the date of recovery to 1980-like conditions varies from about 2045 to 2055.

The results of Figure 8 show similar results for the Spring for 3-D models. However, in the Arctic, most models do not show substantial Arctic ozone change throughout the period 2020 to 2050, while the low values of the GISS model for the decade 2010 to 2020, have disappeared by 2030. In the Antarctic, the recovery of spring ozone, already underway by 2020, continues in the simulations completed (Figure 8). Recovery to 1980-like conditions occurs in the CCSR/NIES and GISS models by about 2045, but perhaps a decade later in the UMetrac snapshot results. The CCSR/NIES and GISS transient model results suggest a near monotonic recovery of ozone, but the UMetrac snapshot results suggest that ozone could undergo further loss over the period 2025 to 2045. This was identified as due to increases in ice PSCs as the lower stratospheric climate cools (Austin et al., 2001), but would need to be confirmed by more model simulations.

5. Conclusions

A number of processes influencing the modeling of polar and mid-latitude ozone have been discussed, including temperature biases, PSCs and the sedimentation of PSC particles, the position of the upper boundary, interannual variability, planetary wave amplitudes and the rate of increase of water vapor.

Cold biases have been found to exist in the stratosphere of many chemistry-climate models, consistent with that previously found for models without chemistry (Pawson et al., 2000). As indicated by Newman et al. (2001), there appears to be a link with the simulation of heat fluxes into the stratosphere. The results here suggest that increased horizontal resolution improves this aspect, while incorporation of non-orographic gravity wave drag can reduce overall temperature biases, as indicated by Manzini and McFarlane (1998). In the lower stratosphere, most models have a cold bias in the spring, when heterogeneous ozone depletion is most likely.

Climate models have been unable to simulate the observed increase in water vapor concentrations, except when there is a significant warming of the tropical
that these may lead to a significant cooling of the polar lower stratosphere, with the potential to enhance PSC formation and hence ozone loss.

The completion of several 3-D simulations of chemistry-climate coupling now provide a better indication of likely future changes than was previously possible. The results suggest that the start of ozone recovery in the Arctic may be delayed by a few years by greenhouse gas increases, but some models give conflicting results. Differences in the simulation of gravity waves and planetary waves as well as model resolution are likely major sources of uncertainty for this issue. However, the processes leading to stratospheric ozone depletion may be too chaotic to provide a definitive answer of how the Arctic will evolve in reality. In the Antarctic, the ozone hole has probably reached its deepest although the vertical and horizontal extent of depletion may increase slightly further over the next few years. Despite this, the total column will likely begin to recover within the next few years, although interannual variability will tend to mask the signal for the next decade. On the longer timescale, to the middle of the 21st century, model predictions appear to be more uncertain. Hence, although recovery of ozone to 1980-like conditions is to be expected by about 2055 (e.g., Rosenfield et al., 2001), models will need to have a better representation of the water vapor increase than has hitherto been possible for full confidence in their predictions.

References

Brühl, C., B. Steil, and E. Manzini, Feedback processes between chemistry and meteorology with focus on lower stratospheric polar vortices, simulations with a coupled GCM, Proceedings SPARC, November 2000, Mar del Plata, Argentina, Available at the SPARC website or CD-ROM), 2000.
Butchart, N., J. Austin, J.R. Knight, A.A. Scaife, and M.L. Gallani, The response of the stratospheric climate to projected changes in the concentrations of the


Groves, K.S. and A.F. Tuck, Stratospheric $O_3$ - $CO_2$ coupling in a photochemical model. I: Without chlorine chemistry, II: With chlorine chemistry, Q.J.R. Mete-


Rind D., D. Shindell, P. Lonergan, and N.K. Balachandran, Climate change and the


Yang, F., M.E. Schlesinger, and E. Rozanov, Description and performance of the UIUC 24-layer stratosphere/troposphere general circulation model, J. Geophys.
Zurek, R. W., G. L. Manney, A. J. Miller, M. E. Gelman, and R. M. Nagatani,
Interannual variability of the north polar vortex in the lower stratosphere during
POPULAR SUMMARY

Over the last decade, improved computer power has allowed 3 dimensional models of the stratosphere to be developed that can be used to simulate polar ozone levels over long periods. This paper compares the meteorology between these models, and discusses the future of polar ozone levels over the next 50 years.